

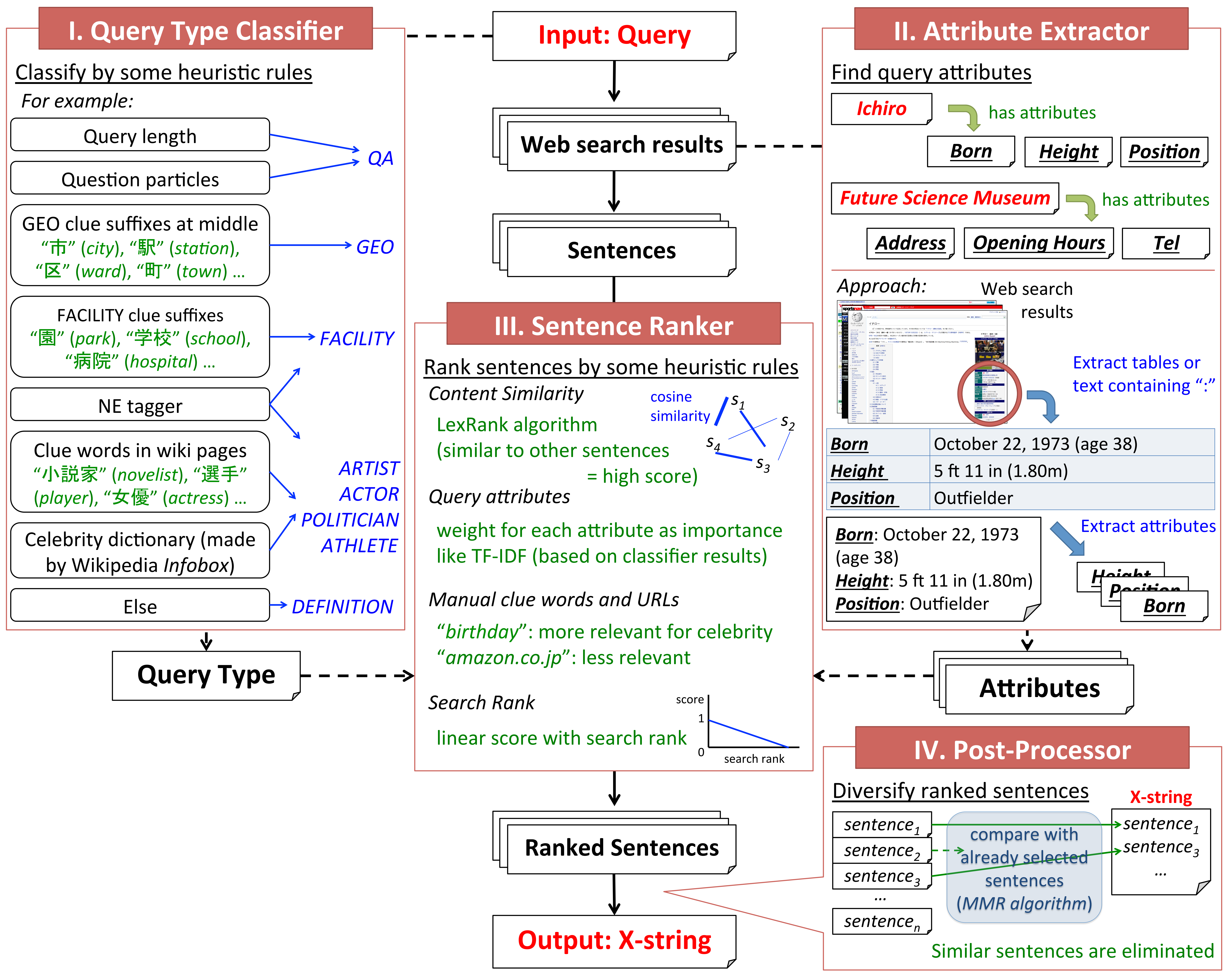
MSRA at NTCIR-10 1CLICK-2

Kazuya Narita (Tohoku University, Japan),
Tetsuya Sakai, Zhicheng Dou (MSRA, China),
Young-In Song (NHN Corporation, Korea)

Takeaway

► Automatic query attribute extraction is effective especially for celebrity queries.

System Architecture



Evaluation Results and Discussion

► Query classification accuracy: 0.83

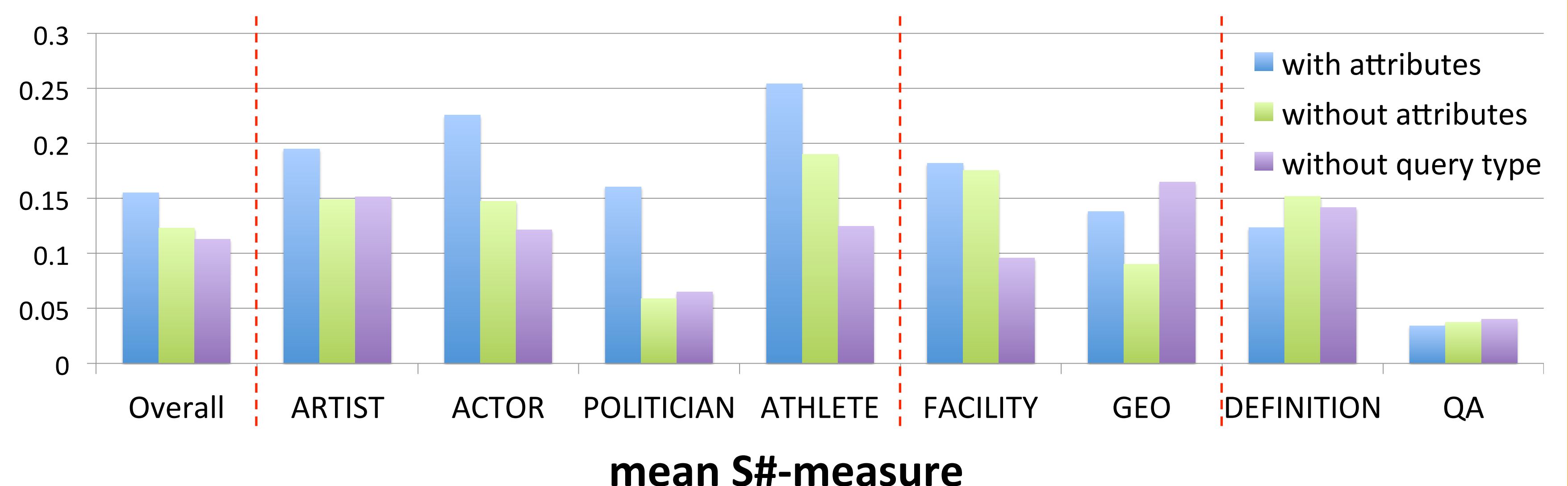
▷ FACILITY: insufficient suffixes and other rules

gold \ sys	ART	ACT	POL	ATH	FAC	GEO	DEF	QA
ARTIST	9	1	0	0	0	0	0	0
ACTOR	1	9	0	0	0	0	0	0
POLITICIAN	1	0	8	0	0	0	1	0
ATHLETE	0	0	0	10	0	0	0	0
FACILITY	1	0	0	0	7	1	5	1
GEO	1	0	0	0	0	14	0	0
DEFINITION	0	0	0	1	1	1	12	0
QA	0	0	0	0	0	0	1	14

► Query attributes are effective especially for celebrity queries

▷ E.g. ATHLETE query “イチロー” (Ichiro; professional baseball player): effective attributes “所属” (affiliation) and “打率” (batting average)

► DEFINITION and QA: attributes do not work well (by the nature of types)



Conclusion and Future work

► Automatic query attribute extraction is effective especially for celebrity queries.

► Future work: automatic extraction of clue words for query classification, new framework instead of attributes for DEFINITION and QA